Tropicana Write Up

MKTG-462 Retail Analytics

Eknoor Sethi

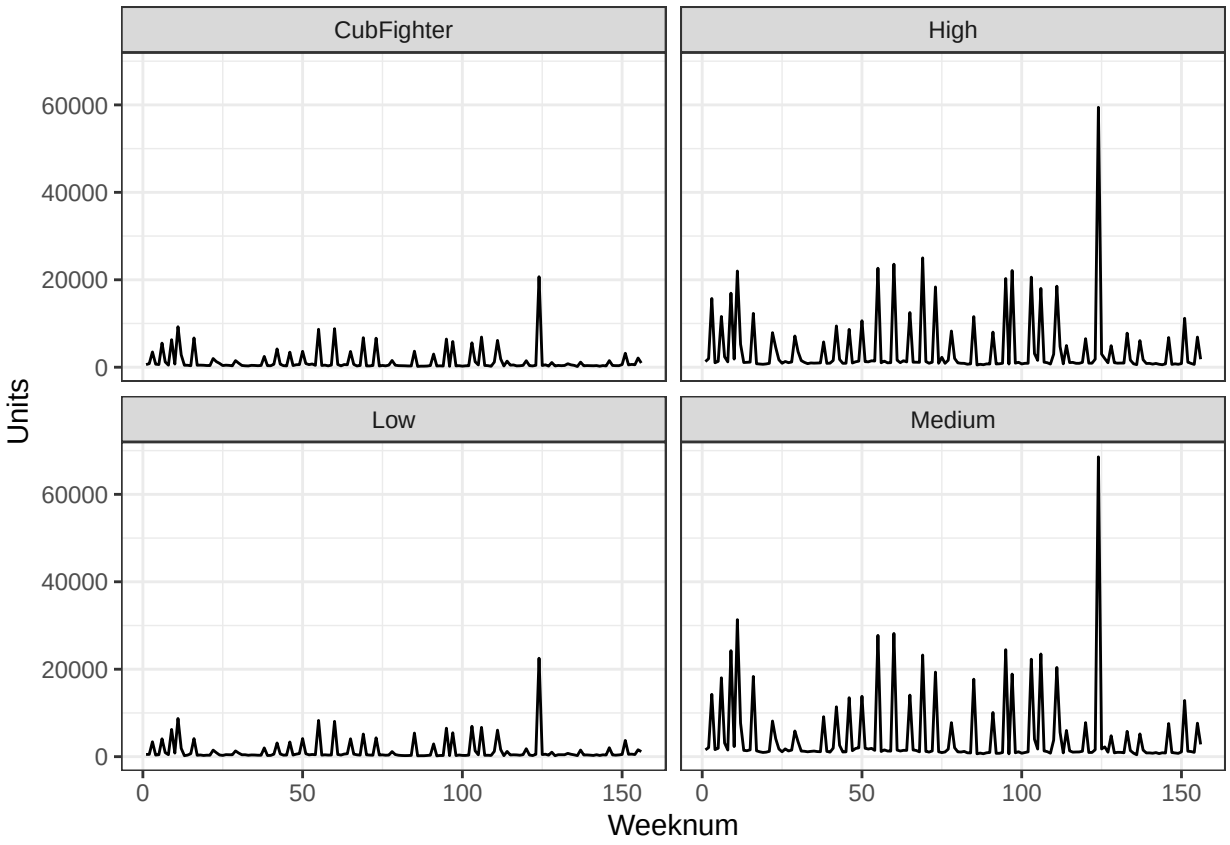
04/28/2025

1 Technical section

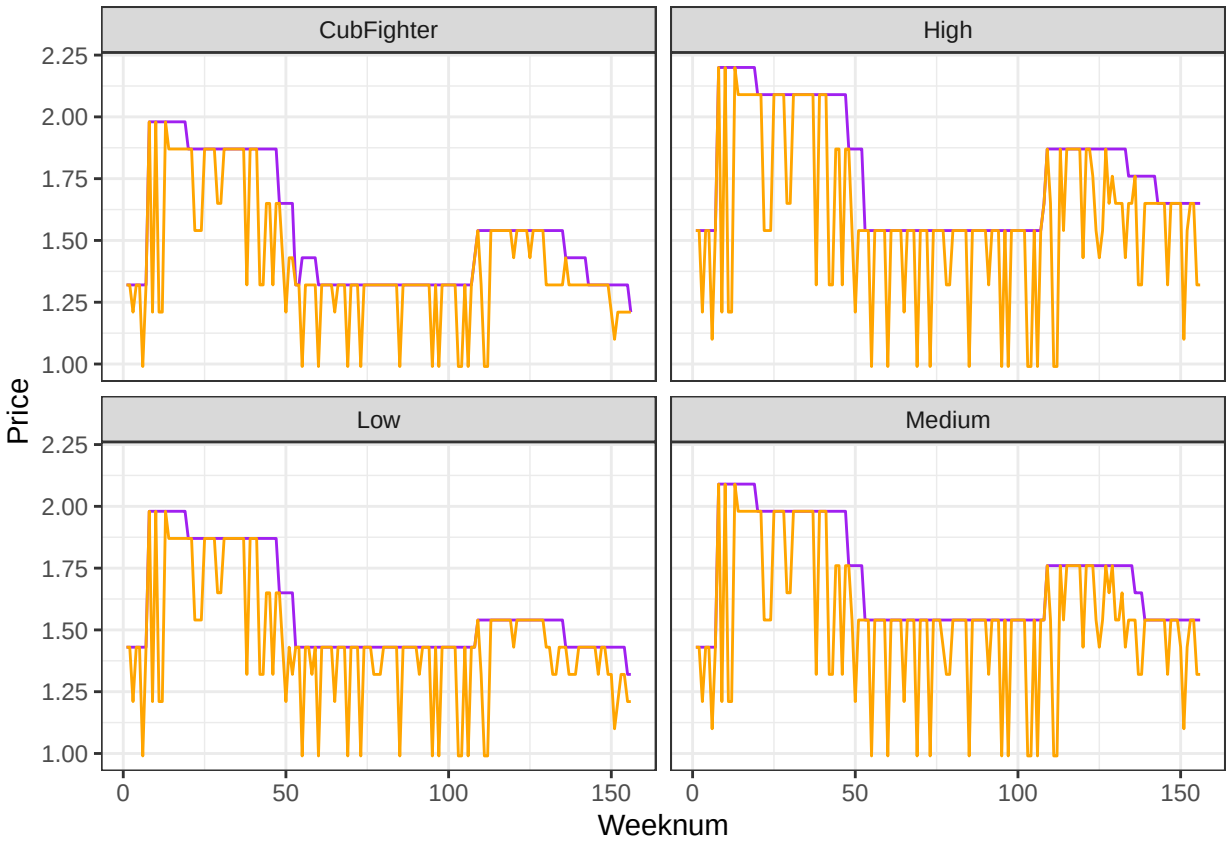
```
# Read the data
df <- read_excel("tropicana.xlsx")

# Create some variables
df <- df %>% mutate(ln_p=log(price_tr_12), ln_q=log(units_tr_12), Dmerch=factor(merch_tr
```

1.1 Explore Your Data

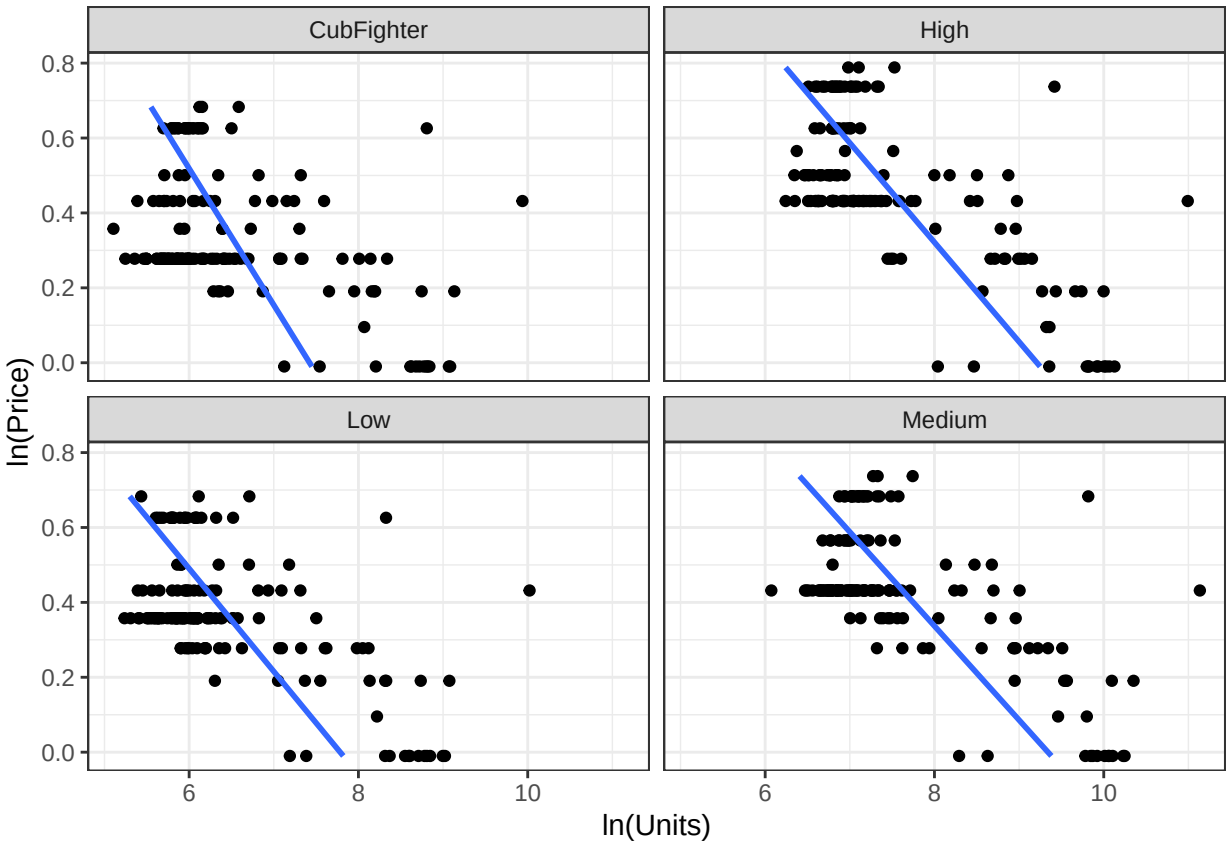
```
# Plot Units Sold
ggplot(df, aes(x=weeknum, y=units_tr_12)) + geom_line() + facet_wrap( ~ zone, ncol=2) +
```



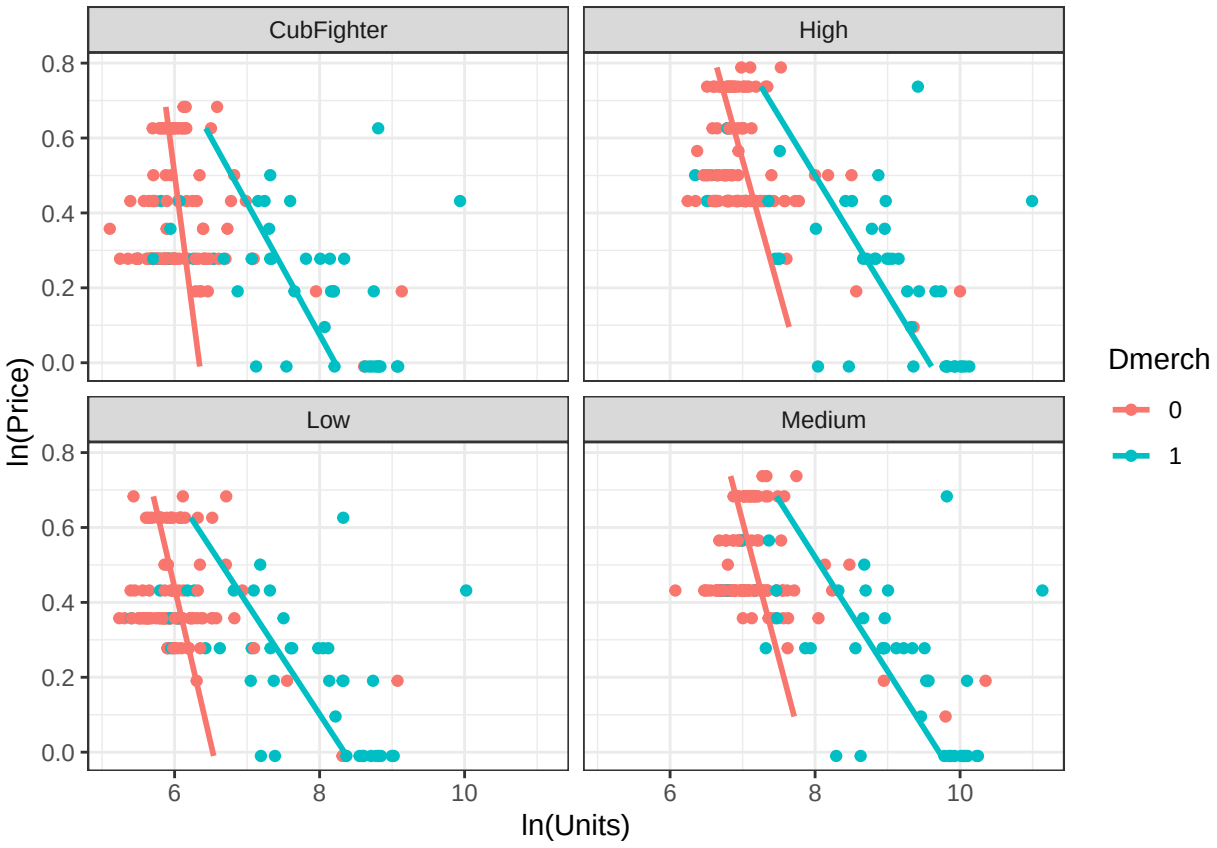

```
# Plot Prices
ggplot(df, aes(x=weeknum)) + geom_line(aes(y=reg_tr_12), color="purple") + geom_line(ae
```




```
# Plot Log Price vs. Log Units
ggplot(df, aes(x=ln_q, y=ln_p)) + geom_point() + facet_wrap( ~ zone, ncol=2) + geom_smo
```




```
ggplot(df, aes(x=ln_q, y=ln_p, col=Dmerch)) + geom_point() + facet_wrap( ~ zone, ncol=2
```



1.2 Run the main regression model

```
reg.tr <- lm(ln_q ~ ln_p + ln_p:Dmerch + Dmerch + Dyear + Dzone, data=df)
summary(reg.tr)
```

```
Call:
lm(formula = ln_q ~ ln_p + ln_p:Dmerch + Dmerch + Dyear + Dzone,
    data = df)

Residuals:
    Min      1Q  Median      3Q     Max
-1.8147 -0.2814 -0.0223  0.2561  3.3869

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.62497    0.12316  61.909  < 2e-16 ***
ln_p         -2.71574    0.22648 -11.991  < 2e-16 ***
Dmerch1       1.38952    0.12162  11.425  < 2e-16 ***
```

```
Dyear2017    -0.82689    0.06810 -12.141  < 2e-16 ***
Dyear2018    -0.78565    0.06209 -12.654  < 2e-16 ***
DzoneHigh     1.33441    0.07070  18.875  < 2e-16 ***
DzoneLow      0.03781    0.06654   0.568 0.570057
DzoneMedium   1.40082    0.06888  20.339  < 2e-16 ***
ln_p:Dmerch1 -1.06353    0.31220  -3.407 0.000701 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.585 on 615 degrees of freedom
Multiple R-squared:  0.7588,    Adjusted R-squared:  0.7557
F-statistic: 241.9 on 8 and 615 DF,  p-value: < 2.2e-16
```

1.3 Scenario Analysis

1.3.1 Scenario 0 (Baseline)

```
# Predict ln(units) using the actual data -- the output are the fitted values
df.pred <- df
df.pred$ln_pred_units_0 <- predict(reg.tr, df.pred)

# Convert Log to Units and Compute Profit
df.pred <- df.pred %>% mutate(units_0 = exp(ln_pred_units_0), profit_0 = units_0 * (pric

# Create a table to summarize predicted profit for baseline
baseline_summary <- df.pred %>% group_by(year) %>% summarize(sum_units_0 = sum(units_0,

print("Summary of Predicted Units and Profit for Baseline:")
```

```
[1] "Summary of Predicted Units and Profit for Baseline:"
```

```
print(baseline_summary)
```

year	sum_units_0	avg_units_per_week_0	sum_profits_0	avg_profit_per_week_0
2016	482284	2319	184180	885
2017	467286	2291	120271	590
2018	411595	1941	134663	635

1.3.2 Scenario 1

```
# Implement the scenario!
# You need to modify the original data to alter the price and merchandise variables
# using the *_scenario variables that we provide.
df_scenario1 <- df %>% filter(year == 2018) %>% mutate(price_tr_12 = price_tr_12_scenari

# Use the regression model to predict using the modified data
df_scenario1$ln_pred_units_scenario1 <- predict(reg.tr, newdata = df_scenario1)

# Convert Log to Units and Compute Profit
df_scenario1$pred_units_scenario1 <- exp(df_scenario1$ln_pred_units_scenario1)

df_scenario1$profit_scenario1 <- (df_scenario1$price_tr_12 - df_scenario1$cost_tr_12) *

# Create a table to summarize predicted profit for the alternative scenario
scenario1_summary <- df_scenario1 %>% summarize(sum_units_1 = sum(pred_units_scenario1,

print("Summary of Predicted Units and Profit for Scenario 1:")
```

```
[1] "Summary of Predicted Units and Profit for Scenario 1:"
```

```
print(scenario1_summary)
```

sum_units_1	avg_units_per_week_1	total_profit_1	avg_profit_per_week_1
290592	1371	112492	531

2 Managerial Discussion

2.1 Overview

As we analyzed the historical situation in Scenario 0 (as the baseline), we can see that profits are still down in 2018 compared to 2016. The profits in 2016 were $184,180, in 2017 were $120,271, and in 2018 were $134,663. Although profits were higher in 2018 compared to 2017, both years were much lower in profits compared to 2016. The analysis and report determines that fewer promotions will not improve the overall profitability of Tropicana.

2.2 Data Exploration: Compare the Scenarios

As previously stated, profits were lower in 2018 as compared to 2016 by $49,517. In Scenario 1, the units sold and total profit was much lower than the baseline Scenario 0 for 2018. In Scenario 1, total units sold were 290,592 and total profits were $112,492, whereas in the baseline Scenario 0, total units sold were 411,595 and total profits were $134,663. This shows us that the predictions for Scenario 1 are much less profitable than the baseline Scenario 0. Overall when comparing the two scenarios, it's clear that fewer promotions will lead to less profitability.

2.3 Managerial Description of Models & Analysis

To determine what the impact of fewer promotions would be on Tropicana, we created two analyses to compare scenarios based on historical data and predicted data, focusing on 2018.

First, we created some plots to visualize different necessary factors for Tropicana. The first plot showed the relationship between the weekly units sold for Tropicana based on which zone it was in. The second plot showed the relationship between pricing strategy and weeks, based on zone as well. The third plot showed the relationship between Log Price and Log Units based on zone. Finally, the fourth plot showed the relationship between price and units sold based on promotions.

Then, a regression model for the historical data from 2016 to 2018 was created to understand the relationship between price, merchandising, and units sold. By using this regression model, we predicted the units sold and calculated profits for the baseline Scenario 0 and promotional Scenario 1. When comparing the two scenarios, as stated previously, it's clear that fewer promotional time leads to lower profitability for Tropicana.

2.4 Managerial Recommendations & Limitations

Based on the data exploration and analysis done, it would be best if Jewel did not go forward with the reduced promotional and merchandising weeks strategy for Tropicana. The data and analysis indicates that merchandising leads to higher units sold and overall profits, so doing the opposite of that will ultimately hurt potential profitability.

There are various limitations within the data and analysis that should be addressed. First, the regression model assumes there is a linear relationship between Log Price and Log Units, and also that the elasticity of demand is not shown to fluctuate depending on the promotional periods. Next, the data and analysis does not take every zone into account, which could be a factor in sales and profit fluctuation. The data and analysis also assumes that there is no fluctuation with the cost of the product, which also goes hand in hand with zone pricing not being taken into account here. Different zones could have different pricing, which could lead to varied sales and profits. Furthermore, the data and analysis assumes that the historical data provided from 2016 until 2018 is reliable in order to predict for the future.